

A Case Study on Logical Relations using Contextual Types

Andrew Cave
McGill University
Montreal QC, Canada
acave1@cs.mcgill.ca

Brigitte Pientka
McGill University
Montreal QC, Canada
bpientka@cs.mcgill.ca

Proofs by logical relations play a key role to establish rich properties such as normalization or contextual equivalence. They are also challenging to mechanize. In this paper, we describe the completeness proof of algorithmic equality for simply typed lambda-terms by Crary where we reason about logically equivalent terms in the proof environment Beluga. There are three key aspects we rely upon: 1) we encode lambda-terms together with their operational semantics and algorithmic equality using higher-order abstract syntax 2) we directly encode the corresponding logical equivalence of well-typed lambda-terms using recursive types and higher-order functions 3) we exploit Beluga’s support for contexts and the equational theory of simultaneous substitutions. This leads to a direct and compact mechanization, demonstrating Beluga’s strength at formalizing logical relations proofs.

1 Introduction

Proofs by logical relations play a fundamental role to establish rich properties such as contextual equivalence or normalization. This proof technique goes back to Tait (26) and was later refined by Girard (12). The central idea of logical relations is to specify relations on well-typed terms via structural induction on the syntax of types instead of directly on the syntax of terms themselves. Thus, for instance, logically related functions take logically related arguments to related results, while logically related pairs consist of components that are related pairwise.

Mechanizing logical relations proofs is challenging: first, specifying logical relations themselves typically requires a logic which allows arbitrary nesting of quantification and implications; second, to establish soundness of a logical relation, one must prove the Fundamental Property which says that any well-typed term under a closing simultaneous substitution is in the relation. This latter part requires some notion of simultaneous substitution together with the appropriate equational theory of composing substitutions. As Altenkirch (1) remarked,

“I discovered that the core part of the proof (here proving lemmas about CR) is fairly straightforward and only requires a good understanding of the paper version. However, in completing the proof I observed that in certain places I had to invest much more work than expected, e.g. proving lemmas about substitution and weakening.”

While logical normalization proofs often are not large, they are conceptually intricate and mechanizing them has become a challenging benchmark for proof environments. There are several key questions, when we attempt to formalize such proofs: How should we represent the abstract syntax tree for lambda-terms and enforce the scope of bound variables? How should we represent well-typed terms or typing derivations? How should we deal with substitution? How can we define the logical relation on closed terms?

Early work (1; 2; 5) represented lambda-terms using (well-scoped) de Bruijn indices which leads to a substantial amount of overhead to prove properties about substitutions such as substitution lemmas

and composition of substitution. To improve readability and generally better support such meta-theoretic reasoning, nominal approaches support α -renaming but substitution and properties about them are specified separately; the Isabelle Nominal package has been used in a variety of logical relations proofs from proving strong normalization for Moggi’s modal lambda-calculus (7) to mechanically verifying the meta-theory of LF itself including the completeness of equivalence checking (16; 27).

Approaches representing lambda-terms using higher-order abstract syntax (HOAS) trees (also called λ -tree syntax) model binders in the object language (i.e. in our case the simply typed lambda-calculus) as binders in the meta language (i.e. in our case the logical framework LF (13)). Such encodings inherit not only α -renaming and substitution from the meta-language, but also weakening and substitution lemmas. However, direct encodings of logical relations proofs is beyond the logical strength supported in systems such as Twelf (17). In this paper, we demonstrate the power and elegance of logical relations proofs within the proof environment Beluga (22) which is built on top of the logical framework LF. Beluga allows programmers to pair LF objects together with their surrounding context and this notion is internalized as a contextual type $[\Psi \vdash A]$ which is inhabited by term M of type A in the context Ψ (15). Proofs about contexts and contextual LF objects are then implemented as dependently-typed recursive functions via pattern matching (18; 21). Beluga’s functional language supports higher-order functions and indexed recursive data-types (3) which we use to encode the logical relation. As such it does not impose any restrictions as for example found in Twelf (17) which does not support arbitrary quantifier alternation or Delphin (23) which lacks recursive data-types. Recently, Beluga has been extended to first-class simultaneous substitutions allowing abstraction over substitutions and supporting a rich equational theory about them (4; 20).

In this paper, we describe the completeness proof of algorithmic equality for simply typed lambda-terms by Crary (6) where we reason about logically equivalent terms in the proof environment Beluga. There are three key aspects we rely upon: 1) we encode lambda-terms together with their operational semantics together with algorithmic equality using higher-order abstract syntax 2) we directly encode the corresponding logical equivalence of well-typed lambda-terms using recursive types and higher-order functions 3) we exploit Beluga’s support for contexts and the equational theory of simultaneous substitutions. This leads to a direct and compact mechanization and allows us to demonstrate Beluga’s strength at formalizing logical relations proofs. Based on this case study we also draw some general lessons.

2 Proof Overview: Completeness of Algorithmic Equality

In this section we give a brief overview of the motivation and high level structure of the completeness proof of algorithmic equality. For more detail, we refer the reader to (6) and (14). Extensions of this proof are important for the metatheory of dependently typed systems such as LF and varieties of Martin-Löf Type Theory, where they are used to establish decidability of typechecking. The proof concerns three judgements, the first of which is declarative equivalence:

$$\Gamma \vdash M \equiv N : A \quad \text{terms } M \text{ and } N \text{ are declaratively equivalent at type } A$$

Declarative equivalence includes convenient but non-syntax directed rules such as transitivity and symmetry, among rules for congruence, extensionality and β -contraction. We will see the full definition in Sec. 3. In particular, it may include apparently type-directed rules such as extensionality at unit type:

$$\frac{\Gamma \vdash M : \text{Unit} \quad \Gamma \vdash N : \text{Unit}}{\Gamma \vdash M \equiv N : \text{Unit}}$$

This rule relies crucially on type information, so the common untyped rewriting strategy for deciding equivalence no longer applies. Instead, one can define an algorithmic notion of equivalence which is directed by the syntax of types. This is the path we follow here. We define algorithmic term equivalence mutually with path equivalence, which is the syntactic equivalence of terms headed by variables, i.e. terms of the form $xM_1 \dots M_n$.

$$\begin{aligned} \Gamma \vdash M \Leftrightarrow N : A & \text{ terms } M \text{ and } N \text{ are algorithmically equivalent at type } A \\ \Gamma \vdash M \leftrightarrow N : A & \text{ paths } M \text{ and } N \text{ are algorithmically equivalent at type } A \end{aligned}$$

In what follows, we sketch the proof of completeness of algorithmic equivalence for declarative equivalence. A direct proof by induction over derivations fails unfortunately in the application case where we need to show that applying equivalent terms to equivalent arguments yields equivalent results, which is not so easy. Instead, one can proceed by proving a more general statement that declaratively equivalent terms are *logically equivalent*, and so in turn algorithmically equivalent. Logical equivalence is a relation defined directly on the structure of the types. We write it as follows:

$$\Gamma \vdash M \approx N : A \text{ Terms } M \text{ and } N \text{ are logically equivalent at type } A$$

The key case is at function type, which directly defines logically equivalent terms at function type as taking logically equivalent arguments to logically equivalent results. Crary defines:

$$\begin{aligned} \Gamma \vdash M_1 \approx M_2 : A \Rightarrow B & \text{ iff } \text{for all } \Delta \geq \Gamma \text{ and } N_1, N_2, \\ & \text{if } \Delta \vdash N_1 \approx N_2 : A \\ & \text{then } \Delta \vdash M_1 N_1 \approx M_2 N_2 : B \end{aligned}$$

A key complication is the quantification over all extensions Δ of the context Γ . This is essential to show completeness of the algorithmic rule for function types, which states that to compare two terms $\Gamma \vdash M \Leftrightarrow N : A \Rightarrow B$ it suffices to compare their applications to *fresh* variables: $\Gamma, x : A \vdash Mx \Leftrightarrow Nx : B$. The generalization to *all* extensions Δ of Γ then arises naturally. This Kripke-style monotonicity condition is one of the reasons that this proof is more challenging than normalization proofs for simply typed lambda-terms, where this quantification can often be avoided using other technical tricks.

For our formalization, we take a slightly different approach which better exploits the features of Beluga available to us. We instead quantify over an arbitrary context Δ together with a simultaneous substitution π which provides for each $x:T$ in Γ , a path M satisfying $\Delta \vdash M \leftrightarrow M : T$. We will call such a substitution a *path substitution* and write this condition as $\Delta \vdash \pi : \Gamma$. In the course of the completeness proof, π will actually only ever be instantiated by substitutions which simply perform weakening. That is, Δ will be of the form Γ, Γ' where $\Gamma = x_1:A_1, \dots, x_n:A_n$ and π will be of the form $\Gamma, \Gamma' \vdash x_1/x_1, \dots, x_n/x_n : \Gamma$. However, the extra generality of path substitutions surprisingly does no harm to the proof, and fits well within Beluga.

$$\begin{aligned} \Gamma \vdash M_1 \approx M_2 : A \Rightarrow B & \text{ iff } \text{for all } \Delta, \text{ path substitutions } \Delta \vdash \pi : \Gamma, \text{ and } N_1, N_2 \\ & \text{if } \Delta \vdash N_1 \approx N_2 : A \\ & \text{then } \Delta \vdash M_1[\pi] N_1 \approx M_2[\pi] N_2 : B \end{aligned}$$

The high level goal is to establish that declaratively equivalent terms are logically equivalent, and that logically equivalent terms are algorithmically equivalent. The proof requires establishing a few key properties of logical equivalence. The first is monotonicity, which is crucially used for weakening logical equivalence. This is used when applying terms to fresh variables.

Lemma 2.1 (Monotonicity)

If $\Gamma \vdash M \approx N : A$ and $\Delta \vdash \pi : \Gamma$ is a path substitution, then $\Delta \vdash M[\pi] \approx N[\pi] : A$

The second key property is (backward) closure of logical equivalence under weak head reduction. This is proved by induction on the type A .

Lemma 2.2 (Logical weak head closure)

If $\Gamma \vdash N_1 \approx N_2 : A$ and $M_1 \longrightarrow_{wh}^* N_1$ and $M_2 \longrightarrow_{wh}^* N_2$ then $\Gamma \vdash M_1 \approx M_2 : A$

In order to escape logical equivalence to obtain algorithmic equivalence in the end, we need the main lemma, which is a mutually inductive proof showing that path equivalence is included in logical equivalence, and logical equivalence is included in algorithmic equivalence:

Lemma 2.3 (Main lemma)

1. If $\Gamma \vdash M \leftrightarrow N : A$ then $\Gamma \vdash M \approx N : A$
2. If $\Gamma \vdash M \approx N : A$ then $\Gamma \vdash M \leftrightarrow N : A$

Also required are symmetry and transitivity of logical equivalence, which in turn require symmetry and transitivity of algorithmic equivalence, determinacy of weak head reduction, and uniqueness of types for path equivalence. We will not go into detail about these lemmas, as they are relatively mundane, but refer the reader to the discussion in (6).

What remains is to show that declarative equivalence implies logical equivalence. This requires a standard technique to generalize the statement to all instantiations of open terms by related substitutions. If σ_1 is of the form $M_1/x_1, \dots, M_n/x_n$ and σ_2 is of the form $N_1/x_1, \dots, N_n/x_n$ and Γ is of the form $x_1:A_1, \dots, x_n:A_n$, we write $\Delta \vdash \sigma_1 \approx \sigma_2 : \Gamma$ to mean that $\Delta \vdash M_i \approx N_i : A_i$ for all i .

Theorem 2.4 (Fundamental theorem)

If $\Gamma \vdash M \equiv N : A$ and $\Delta \vdash \sigma_1 \approx \sigma_2 : \Gamma$ then $\Delta \vdash M[\sigma_1] \approx N[\sigma_2] : A$

The proof goes by induction on the derivation of $\Gamma \vdash M \equiv N : A$. We show one interesting case in order to demonstrate some sources of complexity.

$$\frac{\Gamma, x : A \vdash M_1 \equiv M_2 : B}{\Gamma \vdash \lambda x. M_1 \equiv \lambda x. M_2 : A \Rightarrow B}$$

Proof Case: $\Gamma \vdash \lambda x. M_1 \equiv \lambda x. M_2 : A \Rightarrow B$

1. Suppose we are given Δ' , a path substitution $\Delta' \vdash \pi : \Delta$ and N_1, N_2 with $\Delta' \vdash N_1 \approx N_2 : A$.
2. We have $\Delta' \vdash \sigma_1[\pi] \approx \sigma_2[\pi] : \Gamma$ (by monotonicity)
3. Hence $\Delta' \vdash (\sigma_1[\pi], N_1/x) \approx (\sigma_2[\pi], N_2/x) : \Gamma, x : A$ (by definition)
4. Hence $\Delta' \vdash M_1[\sigma_1[\pi], N_1/x] \approx M_2[\sigma_2[\pi], N_2/x] : B$ (by induction hypothesis)
5. Hence $\Delta' \vdash M_1[\sigma_1[\pi], x/x][N_1/x] \approx M_2[\sigma_2[\pi], x/x][N_2/x] : B$ (by substitution properties)
6. Hence $\Delta' \vdash (\lambda x. M_1[\sigma_1[\pi], x/x]) N_1 \approx (\lambda x. M_2[\sigma_2[\pi], x/x]) N_2 : B$ (by weak head closure)
7. Hence $\Delta' \vdash ((\lambda x. M_1)[\sigma_1])[\pi] N_1 \approx ((\lambda x. M_2)[\sigma_2])[\pi] N_2 : B$ (by substitution properties)
8. Hence $\Delta \vdash (\lambda x. M_1)[\sigma_1] \approx (\lambda x. M_2)[\sigma_2] : A \Rightarrow B$ (by definition of logical equivalence)

We observe that this proof relies heavily on equational properties of substitutions. Some of this complexity appears to be due to our choice of quantifying over substitutions $\Delta \vdash \pi : \Gamma$ instead of extensions $\Delta \geq \Gamma$. However, we would argue that reasoning instead about extensions $\Delta \geq \Gamma$ does not remove this complexity, but only rephrases it.

Finally, by establishing the relatedness of the identity substitution to itself, i.e. $\Gamma \vdash \text{id} \approx \text{id} : \Gamma$ we can combine the fundamental theorem with the main lemma to obtain completeness.

Corollary 2.5 (Completeness) If $\Gamma \vdash M \equiv N : A$ then $\Gamma \vdash M \leftrightarrow N : A$

3 Mechanization

We mechanize the development of the declarative and algorithmic equivalence together with its completeness proof in Beluga, a dependently typed proof language built on top of the logical framework LF. The central idea is to specify lambda-terms, small-step semantics, and type-directed algorithmic equivalence in the logical framework LF. This allows us to model bindings uniformly using the LF function space and obviates the need to model and manage names explicitly. Beluga’s proof language allows programmers to encapsulate LF objects together with their surrounding context as contextual objects and provides support for higher-order functions, indexed recursive types, and pattern matching on contexts and contextual objects. We define logical equivalence and (for technical reasons) declarative equivalence using indexed recursive types. All our proofs will then be implemented as recursive functions using pattern matching and pass the totality checker. The complete source code for our development can be found in the directory `examples/logrel` of the Beluga distribution which is available at <https://github.com/Beluga-lang/Beluga>.

3.1 Encoding lambda-terms, typing and reduction in the logical framework LF

Our proof is about a simply-typed lambda calculus with one base type `i`. Extending the proof to support a unit type and products is straightforward. We describe the types and terms in LF as follows, employing HOAS for the representation of lambda abstraction. That is, we express the body of the lambda expression as an LF function $\text{tm} \rightarrow \text{tm}$. There is no explicit case for variables; they are implicitly handled by LF. We show side by side the corresponding grammar.

<pre> LF tp : type = i : tp => : tp → tp → tp % infix ; LF tm : type = app : tm → tm → tm lam : (tm → tm) → tm; </pre>	<pre> Types T,S ::= i T ⇒ S Terms M,N ::= x lam.x.M app M N </pre>
---	---

Finally, we describe also weak head reduction for our terms. Notice here that the substitution of `N` into `M` in the β -reduction case is accomplished using LF application. We then describe multi-step reductions as a sequence of single step reductions. All free variables occurring in the LF signature are reconstructed and bound implicitly at the outside.

```

LF step : tm → tm → type =
| beta : step (app (lam M) N) (M N)
| stepapp : step M M' → step (app M N) (app M' N);

LF mstep : tm → tm → type =
| refl : mstep M M
| trans1 : step M M' → mstep M' M'' → mstep M M'';

```

3.2 Encoding algorithmic equivalence

We now describe the algorithmic equality of terms. This is defined as two mutually recursive LF specifications. We write `algeqTm M N T` for algorithmic equivalence of terms `M` and `N` at type `T` and `algeqP P Q T` for algorithmic path equivalence at type `T` – these are terms whose head is a variable, not a lambda abstraction. Term equality is directed by the type, while path equality is directed by the syntax. Two terms `M` and `N` at base type `i` are equivalent if they weak head reduce to weak head normal terms `P` and `Q` which are path equivalent. Two terms `M` and `N` are equivalent at type $T \Rightarrow S$ if applying them to a fresh variable `x`

of type τ yields equivalent terms. Variables are only path equivalent to themselves, and applications are path equivalent if the terms at function position are path equivalent, and the terms at argument positions are term equivalent.

```

LF algeqTm: tm → tm → tp → type =
| albase: mstep M P → mstep N Q → algeqP P Q i → algeqTm M N i.
| algarr : ({x:tm} algeqP x x T → algeqTm (app M x) (app N x) S) → algeqTm M N (arr T S)
and algeqP : tm → tm → tp → type
| algapp : algeqP M1 M2 (arr T S) → algeqTm N1 N2 T → algeqP (app M1 N1) (app M2 N2) S;

```

By describing algorithmic equality in LF, we gain structural properties and substitution for free. For this particular proof, only weakening is important.

A handful of different forms of contexts are relevant for this proof. We describe these with **schema** definitions in Beluga. Schemas classify contexts in a similar way as LF types classify LF objects. Although schemas are similar to Twelf's world declarations, schema checking does not involve verifying that a given LF type family only introduces the assumptions specified in the schema; instead schemas will be used by the computation language to guarantee that we are manipulating contexts of a certain shape. Below, we define the schema `actx`, which enforces that term variables come paired with an algorithmic equality assumption `algeqP x x t` for some type `t`.

```

schema actx = some [t:tp] block x:tm, ax:algeqP x x t;

```

3.3 Encoding logical equivalence

To define logical equivalence, we need the notion of path substitution mentioned in Sec. 2. For this purpose, we use Beluga's built-in notion of simultaneous substitutions. We write $[\delta \vdash \gamma]$ for the built-in type of simultaneous substitutions which provide for each variable in the context γ a corresponding term in the context δ . When γ is of schema `actx`, such a substitution consists of blocks of the form $M/x, P/ax$ where M is a term and P is a derivation of `algeqP M M T`, just as we need.

To achieve nice notation, we define an LF type of pairs of terms, where the infix operator \approx simply constructs a pair of terms:

```

LF tmpair : type =
| ≈ : tm → tm → tmpair % infix;

```

Logical equivalence, written $\text{Log } [\gamma \vdash M \approx N] [A]$, expresses that M and N are logically related in context γ at type A . We embed contextual objects into computations and recursive types wrapping them inside $[\]$. Since M and N are used in the context γ , by default, they can depend on γ . Formally, each of these meta-variables is associated with an identity substitution which can be omitted.

We define $\text{Log } [\gamma \vdash M \approx N] [A]$ in Beluga as a *stratified* type, which is a form of recursive type which is defined by structural recursion on one of its indices, as an alternative to an inductive (strictly positive) definition. Beluga verifies that this stratification condition is satisfied. In this case, the definition is structurally recursive on the type A .

```

stratified Log : (γ:actx) [γ ⊢ tmpair] → [tp] → ctype =
| LogBase : [γ ⊢ algeqTm M N i] → Log [γ ⊢ M ≈ N] [i]
| LogArr : ({δ:actx}{π:[δ ⊢ γ]}{N1:[δ ⊢ tm]}{N2:[δ ⊢ tm]})
    Log [δ ⊢ N1 ≈ N2] [T] → Log [δ ⊢ app M1[π] N1 ≈ app M2[π] N2] [S]
    → Log [γ ⊢ M1 ≈ M2] [T ⇒ S];

```

At base type, two terms are logically equivalent if they are algorithmically equivalent. At arrow type we employ the monotonicity condition mentioned in Sec. 2: M_1 is related to M_2 in Γ if, for any context Δ , path substitution $\Delta \vdash \pi : \Gamma$, and N_1, N_2 related in Δ , we have that `app M1[π] N1` is related to `app M2[π] N2`

in Δ . We quantify over $(\gamma:\text{actx})$ in round parentheses, which indicates that it is implicit and recovered during reconstruction. Variables quantified in curly braces such as $\{\delta:\text{actx}\}$ are passed explicitly. As in LF specifications, all free variables occurring in constructor definitions are reconstructed and bound implicitly at the outside. They are passed implicitly and recovered during reconstruction.

Crucially, logical equality is monotonic under path substitutions.

```
rec log_monotone : {δ:actx}{π:[δ ⊢ γ]} Log [γ ⊢ M1 ≈ M2] [A] → Log [δ ⊢ M1[π] ≈ M2[π]] [A]
```

We show below the mechanized proof of this lemma only to illustrate the general structure of Beluga proofs. The proof is simply by case analysis on the logical equivalence. In the base case, we obtain a proof P of $\gamma \vdash \text{algeqTm } M \ N \ i$, which we can weaken for free by simply applying π to P . Here we benefit significantly from Beluga's built-in support for simultaneous substitutions; we gain not just weakening by a single variable for free as we would in Twelf, but arbitrary simultaneous weakening. The proof proceeds in the arrow case by simply composing the two substitutions. We use λ as the introduction form for universal quantifications over metavariables (contextual objects), for which we use uppercase and Greek letters, and fn with lowercase letters for computation-level function types (implications).

```
rec log_monotone : {δ:actx}{π:[δ ⊢ γ]} Log [γ ⊢ M1 ≈ M2] [A] → Log [δ ⊢ M1[π] ≈ M2[π]] [A] =
λ δ, π ↦ fn e ↦ case e of
| LogBase [γ ⊢ P] ↦ LogBase [δ ⊢ P[π]]
| LogArr f ↦ LogArr (λ δ', π' ↦ f [δ'] [δ' ⊢ π[π']])
```

The main lemma is mutually recursive, expressing that path equivalence is included in logical equivalence, and logical equivalence is included in algorithmic term equivalence. This enables “escaping” from the logical relation to obtain an algorithmic equality in the end. They are structurally recursive on the type. Crucially, in the arrow case, `reify` instantiates the path substitution π with a weakening substitution in order to create a fresh variable.

```
rec reflect : {A:[tp]} [γ ⊢ algeqP M1 M2 A] → Log [γ ⊢ M1 ≈ M2] [A]
and reify    : {A:[tp]} Log [γ ⊢ M1 ≈ M2] [A] → [γ ⊢ algeqTm M1 M2 A]
```

We can state weak head closure directly as follows. The proof is structurally recursive on the type, which is implicit.

```
rec closed : [γ ⊢ mstep N1 M1] → [γ ⊢ mstep N2 M2] → Log [γ ⊢ M1 ≈ M2] [T]
→ Log [γ ⊢ N1 ≈ N2] [T]
```

3.4 Encoding declarative equivalence

We now define declarative equality of terms, which includes non-algorithmic rules such as transitivity and symmetry. Declarative equality makes use of a schema which lists only term variables, which we write `ctx`.

```
schema ctx = tm;
```

For technical reasons which we will go into more detail on later, we resort to a different treatment of typing contexts. We explicitly represent typing contexts `dctx` as a list of types, and declarative equality as a computation-level inductive datatype, instead of an LF specification.

```
LF dctx : type =
| nil : dctx
| &   : dctx → tp → dctx % infix ;
```

We describe next the result of looking up the type of a variable x in γ in typing context Γ by its position. If x is the top variable of γ , then its type in Γ is the type of the top variable of Γ . Otherwise, if looking up the type of x in γ yields τ , then looking it up in an extended context also yields τ . Here we

write $[\gamma \vdash \text{tm}]$ for the contextual type of terms of type tm in context γ , and $[\text{tp}]$ for (closed) types. We use $\#p$ for a meta-variable standing for an object-level variable from γ (as opposed to a general term).

inductive $\text{Lookup} : \{\Gamma : [\text{dctx}]\}(\gamma : \text{ctx}) [\gamma \vdash \text{tm}] \rightarrow [\text{tp}] \rightarrow \text{ctype} =$
 | $\text{Top} : \text{Lookup} [\Gamma \ \& \ T] [\gamma, x : \text{tm} \vdash x] [T]$
 | $\text{Pop} : \text{Lookup} [\Gamma] [\gamma \vdash \#p] [T] \rightarrow \text{Lookup} [\Gamma \ \& \ S] [\gamma, x : \text{tm} \vdash \#p] [T];$

We write $\text{Decl} [\Gamma] [\gamma \vdash M \approx N] [T]$ for declarative equivalence of M and N at type T . We employ the convention that Γ and Δ stand for typing contexts (of type $[\text{dctx}]$), while γ and δ stand for corresponding term contexts.

inductive $\text{Decl} : \{\Gamma : [\text{dctx}]\}(\gamma : \text{ctx}) [\gamma \vdash \text{tm pair}] \rightarrow [\text{tp}] \rightarrow \text{ctype} =$
 | $\text{DecBeta} : \text{Decl} [\Gamma \ \& \ T] [\gamma, x : \text{tm} \vdash M2 \approx N2] [S] \rightarrow \text{Decl} [\Gamma] [\gamma \vdash M1 \approx N1] [T]$
 $\rightarrow \text{Decl} [\Gamma] [\gamma \vdash \text{app} (\text{lam} (\backslash x. M2)) M1 \approx N2[\dots, N1]] [S]$
 | $\text{Declam} : \text{Decl} [\Gamma \ \& \ T] [\gamma, x : \text{tm} \vdash M \approx N] [S]$
 $\rightarrow \text{Decl} [\Gamma] [\gamma \vdash \text{lam} (\backslash x. M) \approx \text{lam} (\backslash x. N)] [T \Rightarrow S]$
 | $\text{DecExt} : \text{Decl} [\Gamma \ \& \ T] [\gamma, x : \text{tm} \vdash \text{app} M x \approx \text{app} N x] [S]$
 $\rightarrow \text{Decl} [\Gamma] [\gamma \vdash M \approx N] [T \Rightarrow S]$
 | $\text{DecVar} : \text{Lookup} [\Gamma] [\gamma \vdash \#p] [T] \rightarrow \text{Decl} [\Gamma] [\gamma \vdash \#p \approx \#p] [T]$
 | $\text{DecApp} : \text{Decl} [\Gamma] [\gamma \vdash M1 \approx M2] [T \Rightarrow S] \rightarrow \text{Decl} [\Gamma] [\gamma \vdash N1 \approx N2] [T]$
 $\rightarrow \text{Decl} [\Gamma] [\gamma \vdash \text{app} M1 N1 \approx \text{app} M2 N2] [S]$
 | $\text{DecSym} : \text{Decl} [\Gamma] [\gamma \vdash M \approx N] [T] \rightarrow \text{Decl} [\Gamma] [\gamma \vdash N \approx M] [T]$
 | $\text{DecTrans} : \text{Decl} [\Gamma] [\gamma \vdash M \approx N] [T] \rightarrow \text{Decl} [\Gamma] [\gamma \vdash N \approx O] [T]$
 $\rightarrow \text{Decl} [\Gamma] [\gamma \vdash M \approx O] [T];$

Declarative equality includes a β rule, as well as an extensionality rule, which states that for two terms M and N to be equal at type $T \Rightarrow S$, it suffices for them to be equal when applied to a fresh variable of type T . We again remind the reader that all meta-variables are silently associated with the identity substitution; in particular in $[\gamma \vdash \text{lam} (\backslash x. M) \approx \text{lam} (\backslash x. N)]$, the meta-variables M and N are associated with the identity substitution on the context $\gamma, x : \text{tm}$. Note that every meta-variable is associated with a simultaneous substitutions in Beluga. If this substitution is the identity, then it can be omitted. Hence, $[\gamma \vdash \text{lam} (\backslash x. M) \approx \text{lam} (\backslash x. N)]$ is equivalent to writing $[\gamma \vdash \text{lam} (\backslash x. M[\dots, x]) \approx \text{lam} (\backslash x. N[\dots, x])]$. Written in η -contracted form this is equivalent to: $[\gamma \vdash \text{lam} M \approx \text{lam} N]$ or making the identity substitution explicit $[\gamma \vdash \text{lam} M[\dots] \approx \text{lam} N[\dots]]$.

Note that meta-variables associated with simultaneous substitutions do not exist other systems. For example in LF and its implementation in Twelf (17) the context of assumptions is ambient and we cannot express dependencies of LF-variables on them. In Twelf, writing $\text{lam } M$ is equivalent to its η -expanded form $\text{lam } \backslash x. M x$.

3.5 Fundamental theorem

The fundamental theorem requires us to speak of all instantiations of open terms by related substitutions. We express here the notion of related substitutions using inductive types. Trivially, empty substitutions, written as \cdot , are related at empty domain. If σ_1 and σ_2 are related at Γ and M_1 and M_2 are related at T , then σ_1, M_1 and σ_2, M_2 are related at $\Gamma \ \& \ T$. The technical reason we use the schema ctx of term assumptions only is that we would like the substitutions σ_1 and σ_2 to carry only terms M , but *not* derivations $\text{algeqP } M \ M \ T$ (or declarative equality assumptions). If we had used the schema actx or a schema with declarative equality assumptions, the proof of the fundamental theorem would be obligated to construct these derivations, which would be more cumbersome.

inductive $\text{LogSub} : (\gamma : \text{ctx})(\delta : \text{actx})\{\sigma_1 : [\delta \vdash \gamma]\}\{\sigma_2 : [\delta \vdash \gamma]\}\{\Gamma : [\text{dctx}]\} \text{ctype} =$
 | $\text{Nil} : \text{LogSub} [\delta \vdash \cdot] [\delta \vdash \cdot] [\text{nil}]$


```
| Dot : LogSub [h ⊢ σ1] [h ⊢ σ2] [Γ] → Log [δ ⊢ M1 ≈ M2] [T]
      → LogSub [δ ⊢ σ1, M1] [δ ⊢ σ2, M2] [Γ & T]
```

We have a monotonicity lemma for logically equivalent substitutions which is similar to the monotonicity lemma for logically equivalent terms:

```
rec wknLogSub : {π : [δ' ⊢ δ]} LogSub [δ ⊢ σ1] [δ ⊢ σ2] [Γ]
      → LogSub [δ' ⊢ σ1[π]] [δ' ⊢ σ2[π]] [Γ]
```

The fundamental theorem requires a proof that M_1 and M_2 are declaratively equal, together with logically related substitutions σ_1 and σ_2 , and produces a proof that $M_1[\sigma_1]$ and $M_2[\sigma_2]$ are logically related. In the transitivity and symmetry cases, we appeal to transitivity and symmetry of logical equivalence, the proofs of which can be found in the accompanying Beluga code.

```
rec thm : Decl [Γ] [γ ⊢ M1 ≈ M2] [T]
      → LogSub [δ ⊢ σ1] [δ ⊢ σ2] [Γ]
      → Log [δ ⊢ M1[σ1] ≈ M2[σ2]] [T] =
```

We show the `lam` case of the proof term only to make a high-level comparison to the hand-written proof in Sec. 2. Below, one can see that we appeal to monotonicity (`wknLogSub`), weak head closure (`closed`), and the induction hypothesis on the subderivation `d1`. However, remarkably, there is no explicit equational reasoning about substitutions, since applications of substitutions are automatically simplified. We refer the reader to (4) for the technical details of this simplification.

```
fn d, s ↦ case d of
| Declam d1 ↦
  LogArr (λ δ', π, N1, N2 ↦ fn rn ↦
    let ih = thm d1 (Dot (wknLogSub [δ'] [δ] [δ' ⊢ π] s) rn) in
    closed [δ' ⊢ trans1 beta refl] [δ' ⊢ trans1 beta refl] ih
  )
...

```

Completeness is a corollary of the fundamental theorem. Our statement of the completeness theorem is slightly complicated by the fact that declarative equality and algorithmic equality live in different context schemas. To overcome this, we describe a predicate `EmbedSub [Γ] [γ] [γ' ⊢ ι]` which states that ι is a simple weakening substitution which performs the work of moving from term context γ :`ctx` to the corresponding (larger) algorithmic equality context γ' :`actx` with added algorithmic equality assumptions at the types listed in Γ :`[dctx]`. Morally, this ι substitution plays the role of the identity substitution mentioned in Sec. 2.

```
inductive EmbedSub : {Γ : [dctx]} {γ : ctx} {γ' : actx} {ι : [γ' ⊢ γ]} ctype =
| INil : EmbedSub [nil] [] []
| ISnoc : EmbedSub [Γ] [γ] [γ' ⊢ ι]
      → EmbedSub [Γ & T] [γ, x : tm] [γ', b : block x : tm, ax : algeqP x x T ⊢ ι, b.1]
```

It is then straightforward to show that embedding substitutions ι are logically related to themselves using the main lemma.

```
rec embed_log : EmbedSub [Γ] [γ] [γ' ⊢ ι] → LogSub [Γ] [γ] [γ' ⊢ ι] [γ' ⊢ ι]
```

The completeness theorem is stated below, and follows trivially by composing the fundamental theorem with `embed_log` and the main lemma to escape the logical relation.

```
rec completeness : EmbedSub [Γ] [γ] [γ' ⊢ ι] → Decl [Γ] [γ ⊢ M1 ≈ M2] [T]
      → [γ' ⊢ algeqTm M1[ι] M2[ι] T]
```

It is unfortunate that this transportation from γ to γ' is required by the current framework of contextual types, since intuitively the algorithmic equality assumptions in γ' are completely irrelevant for the terms M_1 and M_2 . It's an open problem how to improve on this.

3.6 Remarks

The proof passes Beluga’s typechecking and totality checking. As part of the totality checker, Beluga performs a strict positivity check for inductive types (19; 20), and a stratification check for logical relation-style definitions.

Beluga’s built-in support for simultaneous substitutions is a big win for this proof. The proof of the monotonicity lemma is very simple, since the (simultaneous) weakening of algorithmic equality comes for free, and there is no need for explicit reasoning about substitution equations in the fundamental theorem or elsewhere. We also found that the technique of quantifying over path substitutions as opposed to quantifying over all extensions of a context to work surprisingly well. However, it seems to be non-obvious when this technique will work. In an earlier version of this proof, we had resorted to explicitly enforcing that the substitution π contained only *variables*, limiting its capabilities to weakening, exchange, and contraction. This was done with an inductive datatype like the following, where the contextual type $\#[\delta \vdash \text{tm}]$ contains only *variables* of type tm :

```
datatype IsRenaming : { $\gamma$ :ctx}( $\delta$ :ctx) { $\pi$ : $[\delta \vdash \gamma]$ } ctype =
| Nil : IsRenaming [] [ $\delta \vdash \cdot$ ]
| Cons : { $\#p$ : $\#[\delta \vdash \text{tm}]$ } IsRenaming [ $\gamma$ ] [ $\delta \vdash \pi$ ]  $\rightarrow$  IsRenaming [ $\gamma, x:\text{tm}$ ] [ $\delta \vdash \pi, \#p$ ]
```

We were surprised to learn that in fact this restriction was unnecessary, and we could instead simply directly quantify over path substitutions, as the schema actx we rely on in our proof already effectively restricts the substitutions we can build. However, we suspect that the technique of explicitly restricting to renaming substitutions may still be necessary in some cases, and that it might be convenient to have a built-in type of these renaming-only substitutions.

We remark that the completeness theorem can in fact be executed, viewing it as an algorithm for normalizing derivations in the declarative system to derivations in the algorithmic system. The extension to a proof of decidability would be a correct-by-construction functional algorithm for the decision problem. This is a unique feature of carrying out the proof in a type-theoretic setting like Beluga, where the proof language also serves as a computation language.

Some aspects of this proof could still be improved. In particular, our treatment of the different context schemas and the relationship between them seems unsatisfactory. We had to do a bit of work in order to move terms from $\gamma:\text{ctx}$ to $\gamma':\text{actx}$, and this polluted the final statement of the completeness theorem. It can also be difficult to know when to resort to using an explicit context and a computation-level datatype, like we did for declarative equality. This suggests there is room for improvement in Beluga’s treatment of contexts, and we are exploring possible approaches.

Furthermore, one might argue that having to explicitly apply the path substitutions π to terms like $M[\pi]$ is somewhat unsatisfactory, so one might wish for the ability to directly perform the bounded quantification $\forall \Delta \geq \Gamma$ and a notion of subtyping which permits for example $[\Gamma \vdash \text{tm}] \leq [\Delta \vdash \text{tm}]$. This is also a possibility we are exploring.

Overall, we found that the tools provided by Beluga, especially its support for simultaneous substitutions, worked remarkably well to express this proof and to obviate the need for bureaucratic lemmas about substitutions and contexts, and we are optimistic that these techniques can scale to many other varieties of logical relations proofs.

4 Related Work

Mechanizing proofs by logical relations is an excellent benchmark to evaluate the power and elegance of a given proof development. Because it requires nested quantification and recursive definitions, encoding

logical relations has been particularly challenging for systems supporting HOAS encodings.

There are two main approaches to support reasoning about HOAS encodings: 1) In the proof-theoretic approaches, we adopt a two-level system where we implement a specification logic (similar to LF) inside a higher-order logic supporting (co)inductive definitions, the approach taken in Abella (9), or type theory, the approach taken in Hybrid (8). To distinguish in the proof theory between quantification over variables and quantification over terms, (10) introduce a new quantifier, ∇ , to describe nominal abstraction logically. To encode logical relations one uses recursive definitions which are part of the reasoning logic (11). Induction in these systems is typically supported by reasoning about the height of a proof tree; this reduces reasoning to induction over natural numbers, although much of this complexity can be hidden in Abella. Compared to our development in Beluga, Abella lacks support for modelling a context of assumptions and simultaneous substitutions. As a consequence, some of the tedious basic infrastructure to reason about open and closed terms and substitutions still needs to be built and maintained. Moreover, Abella’s inductive proofs cannot be executed and do not yield a program for normalizing derivations. It is also not clear what is the most effective way to perform the quantification over all *extensions* of a context in Abella.

2) The type-theoretic approaches fall into two categories: we either remain within the logical framework and encode proofs as relations as advocated in Twelf (17) or we build a dependently typed functional language on top of LF to support reasoning about LF specifications as done in Beluga. The former approach lacks logical strength; the function space in LF is “weak” and only represents binding structures instead of computations. To circumvent these limitations, (25) proposes to implement a reasoning logic within LF and then use it to encode logical relation arguments. This approach scales to richer calculi (24) and avoids reasoning about contexts, open terms and simultaneous substitutions explicitly. However, one might argue that it not only requires additional work to build up a reasoning logic within LF and prove its consistency, but is also conceptually different from what one is used to from on-paper proofs. It is also less clear whether the approach scales easily to proving completeness of algorithmic equality due to the need to talk about context extensions in the definition of logical equivalence of terms of function type.

Outside the world of HOAS, (16) have carried out essentially the same proof in Nominal Isabelle, and later (27) tackle the extension from the simply-typed lambda calculus to LF. Relative to their approach, Beluga gains substitution for free, but more importantly, equations on substitutions are silently discharged by Beluga’s built-in support for their equational theory, so they do not even appear in proofs. In contrast, proving these equations manually requires roughly a dozen intricate lemmas.

5 Conclusion

Our implementation of completeness of algorithmic equality takes advantage of key infrastructure provided by Beluga: it utilizes first-class simultaneous substitutions, contexts, contextual objects and the power of recursive types. This yields a direct and compact implementation of all the necessary proofs which directly correspond to their on-paper developments. Moreover, our proof yields an executable program. While more work on Beluga’s frontend will improve and make simpler such developments, we have demonstrated that the core language is not only suitable for standard structural induction proofs such as type safety, but also proofs by logical relations.

References

- [1] Thorsten Altenkirch (1993): *A Formalization of the Strong Normalization Proof for System F in LEGO*. In Marc Bezem & Jan Friso Groote, editors: *International Conference on Typed Lambda Calculi and Applications (TLCA '93)*, *Lecture Notes in Computer Science* 664, Springer, pp. 13–28, doi:10.1007/BFb0037095.
- [2] Stefano Berardi (1990): *Girard Normalization Proof in LEGO*. In: *Proceedings of the First Workshop on Logical Frameworks*, pp. 67–78.
- [3] Andrew Cave & Brigitte Pientka (2012): *Programming with binders and indexed data-types*. In: *39th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL'12)*, ACM Press, pp. 413–424, doi:10.1145/2103656.2103705.
- [4] Andrew Cave & Brigitte Pientka (2013): *First-class substitutions in contextual type theory*. In: *Proceedings of the Eighth ACM SIGPLAN International Workshop on Logical Frameworks and Meta-Languages: Theory and Practice (LFMTP'13)*, ACM Press, pp. 15–24, doi:10.1145/2503887.2503889.
- [5] Catarina Coquand (1992): *A proof of normalization for simply typed lambda calculus writing in ALF*. In: *Informal Proceedings of Workshop on Types for Proofs and Programs*, Dept. of Computing Science, Chalmers Univ. of Technology and Göteborg Univ., pp. 80–87.
- [6] Karl Crary (2005): *Logical Relations and a Case Study in Equivalence Checking*. In Benjamin C. Pierce, editor: *Advanced Topics in Types and Programming Languages*, The MIT Press.
- [7] Christian Doczkal & Jan Schwinghammer (2009): *Formalizing a Strong Normalization Proof for Moggi's Computational Metalanguage: A Case Study in Isabelle/HOL-nominal*. In: *Proceedings of the Fourth International Workshop on Logical Frameworks and Meta-Languages: Theory and Practice (LFMTP'09)*, ACM, pp. 57–63, doi:10.1145/1577824.1577834.
- [8] Amy Felty & Alberto Momigliano (2012): *Hybrid - A Definitional Two-Level Approach to Reasoning with Higher-Order Abstract Syntax*. *J. Autom. Reasoning* 48(1), pp. 43–105, doi:10.1007/s10817-010-9194-x.
- [9] Andrew Gacek (2008): *The Abella Interactive Theorem Prover (System Description)*. In: *4th International Joint Conference on Automated Reasoning*, *Lecture Notes in Artificial Intelligence* 5195, Springer, pp. 154–161, doi:10.1007/978-3-540-71070-7_13.
- [10] Andrew Gacek, Dale Miller & Gopalan Nadathur (2008): *Combining generic judgments with recursive definitions*. In F. Pfenning, editor: *23rd Symposium on Logic in Computer Science*, IEEE Computer Society Press, pp. 33–44, doi:10.1109/LICS.2008.33.
- [11] Andrew Gacek, Dale Miller & Gopalan Nadathur (2009): *Reasoning in Abella about Structural Operational Semantics Specifications*. In: *Proceedings of the International Workshop on Logical Frameworks and Metalanguages: Theory and Practice (LFMTP 2008)*, *Electronic Notes in Theoretical Computer Science (ENTCS)* 228, Elsevier, pp. 85 – 100, doi:10.1016/j.entcs.2008.12.118.
- [12] J.-Y. Girard, Y. Lafont & P. Taylor (1990): *Proofs and types*. Cambridge University Press.
- [13] Robert Harper, Furio Honsell & Gordon Plotkin (1993): *A Framework for Defining Logics*. *Journal of the ACM* 40(1), pp. 143–184, doi:10.1145/138027.138060.
- [14] Robert Harper & Frank Pfenning (2005): *On Equivalence and Canonical Forms in the LF Type Theory*. *ACM Transactions on Computational Logic* 6(1), pp. 61–101, doi:10.1145/1042038.1042041.

- [15] Aleksandar Nanevski, Frank Pfenning & Brigitte Pientka (2008): *Contextual modal type theory*. *ACM Transactions on Computational Logic* 9(3), pp. 1–49, doi:10.1145/1352582.1352591.
- [16] Julien Narboux & Christian Urban (2008): *Formalising in Nominal Isabelle Crary’s Completeness Proof for Equivalence Checking*. *Electr. Notes Theor. Comput. Sci.* 196, pp. 3–18, doi:10.1016/j.entcs.2007.09.014.
- [17] Frank Pfenning & Carsten Schürmann (1999): *System Description: Twelf — A Meta-Logical Framework for Deductive Systems*. In H. Ganzinger, editor: *16th International Conference on Automated Deduction (CADE-16)*, Lecture Notes in Artificial Intelligence (LNAI 1632), Springer, pp. 202–206, doi:10.1007/3-540-48660-7_14.
- [18] Brigitte Pientka (2008): *A type-theoretic foundation for programming with higher-order abstract syntax and first-class substitutions*. In: *35th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL’08)*, ACM Press, pp. 371–382, doi:10.1145/1328438.1328483.
- [19] Brigitte Pientka & Andreas Abel (2015): *Structural Recursion over Contextual Objects*. In Thorsten Altenkirch, editor: *13th International Conference on Typed Lambda Calculi and Applications (TLCA’15)*, Leibniz International Proceedings in Informatics (LIPIcs) of Schloss Dagstuhl, pp. 273–287, doi:10.4230/LIPIcs.TLCA.2015.273.
- [20] Brigitte Pientka & Andrew Cave (2015): *Inductive Beluga: Programming Proofs (System description)*. In: *25th International Conference on Automated Deduction (CADE-25)*, Springer.
- [21] Brigitte Pientka & Joshua Dunfield (2008): *Programming with proofs and explicit contexts*. In: *ACM SIGPLAN Symposium on Principles and Practice of Declarative Programming (PPDP’08)*, ACM Press, pp. 163–173, doi:10.1145/1389449.1389469.
- [22] Brigitte Pientka & Joshua Dunfield (2010): *Beluga: a Framework for Programming and Reasoning with Deductive Systems (System Description)*. In Jürgen Giesl & Reiner Haehnle, editors: *5th International Joint Conference on Automated Reasoning (IJCAR’10)*, Lecture Notes in Artificial Intelligence (LNAI 6173), Springer-Verlag, pp. 15–21, doi:10.1007/978-3-642-14203-1_2.
- [23] Adam B. Poswolsky & Carsten Schürmann (2008): *Practical programming with higher-order encodings and dependent types*. In: *17th European Symposium on Programming (ESOP ’08)*, 4960, Springer, pp. 93–107, doi:10.1007/978-3-540-78739-6_7.
- [24] Ulrik Rasmussen & Andrzej Filinski (2013): *Structural Logical Relations with Case Analysis and Equality Reasoning*. In: *Proceedings of the Eighth ACM SIGPLAN International Workshop on Logical Frameworks and Meta-languages: Theory and Practice (LFMTP’13)*, ACM Press, pp. 43–54, doi:10.1145/2503887.2503891.
- [25] Carsten Schürmann & Jeffrey Sarnat (2008): *Structural Logical Relations*. In: *23rd Annual Symposium on Logic in Computer Science (LICS)*, Pittsburgh, PA, USA, IEEE Computer Society, pp. 69–80, doi:10.1109/LICS.2008.44.
- [26] William Tait (1967): *Intensional Interpretations of Functionals of Finite Type I*. *J. Symb. Log.* 32(2), pp. 198–212, doi:10.2307/2271658.
- [27] Christian Urban, James Cheney & Stefan Berghofer (2011): *Mechanizing the metatheory of LF*. *ACM Trans. Comput. Log.* 12(2), p. 15, doi:10.1145/1877714.1877721.